

Characterization of small RNAs in microsporidian *Nosema bombycis*

PAN Qiu-Ling¹, LI Tian², HE Qiang¹, MA Zhen-Gang¹, FAN Xiao-Dong¹,
ZHANG Xiao-Yan¹, WANG Yan-Li¹, ZHOU Ze-Yang^{1,2}, XU Jin-Shan^{1,*}

(1. College of Life Sciences, Chongqing Normal University, Chongqing 401331, China;

2. State Key Laboratory of Silkworm Genome Biology, Southwest University, Chongqing 400716, China)

Abstract: 【Aim】 The microsporidian *Nosema bombycis* can produce epizootic outbreaks in populations of the domestic silkworm (*Bombyx mori*) by the major mode of vertical transmission and horizontal transmission. In order to explore the genomic defense system against repetitive elements and potential means of gene regulation in *N. bombycis*, we conducted the survey of small RNAs associated with transposons and identification of the potential miRNAs in this silkworm parasite. 【Methods】 Total RNAs were extracted from silkworm midguts infected by *N. bombycis* spores and small RNA cDNA libraries were constructed for deep sequencing. After Solexa sequencing, we performed the functional characterization of *N. bombycis* small RNAs based on the reference genomic sequence through bioinformatics. 【Results】 *N. bombycis* has predominantly small RNAs with the length of 24 and 25 nt, most of which show a strong base bias for U at the 5' end. Redundant small RNAs were associated with transposons in *N. bombycis*, and the amount of antisense small RNAs aligned with transposons was much more than that of sense small RNAs. Thirty-one candidate miRNAs were characterized by means of computational prediction, and certain miRNAs were shared by other *Nosema* species, suggesting that they may evolve conservatively. 【Conclusion】 There should exist small RNAs genomic defense system against transposons in *N. bombycis*. Tentative identification of potential thirty-one miRNAs provides a basis for further functional research of their targets.

Key words: *Nosema bombycis*; small RNAs; transposon; miRNA; genome

1 INTRODUCTION

Since small RNAs play important roles in the development of individual organisms and regulation (Kawaoka *et al.*, 2008). In recent decades, numerous studies have focused on small non-coding RNAs from various organisms. According to the mechanism of their biosynthesis pathways, small RNAs can be divided into three different categories: microRNAs (miRNAs), small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs) (Lau *et al.*, 2009). siRNAs were generated as 21–25 nt in length through a long double-stranded RNA molecule digested by a Dicer enzyme, and then silenced the transcription of targeted genes by binding with AGO proteins (Ender *et al.*, 2008; Höck and Meister, 2008; Obbard and Finnegan, 2008). Mature miRNAs, about 19–25 nt in length, were generated from the precursor miRNAs (primary miRNAs) containing hairpin structure through the action of the RNase III family enzymes

Drosha and Dicer, and were then bound by Ago-subfamily proteins (Sinkkonen *et al.*, 2008; Kim *et al.*, 2009). miRNA regulated gene expression at the post-transcriptional level, and mediated several important biochemical pathways (MacRae *et al.*, 2007). piRNAs, initially known as rasiRNAs, were a class of newly discovered small RNAs with the length of 26–32 nt (Girard *et al.*, 2006; Houwing *et al.*, 2007; Wang *et al.*, 2014). Generation of piRNAs depended on the role of PIWI subfamily proteins, but not Dicer enzyme (Vagin *et al.*, 2006). In *Drosophila*, piRNAs were initially named as repeat-associated small interfering RNAs (rasiRNAs), due to mainly derivation from the repetitive DNA sequences (Saito *et al.*, 2006; Brennecke *et al.*, 2007).

The function of small RNAs had been reported in several fungi and protists. For instance, small RNAs of *Candida albicans* were able to silence endogenous retrotransposon activity (Drinnenberg *et al.*, 2009); small RNAs from *Schizosaccharomyces*

基金项目: 国家高技术研究发展计划项目(2013AA102507); 国家自然科学基金项目(31001037, 31270138, 31402142); 国家级大学生创新创业训练计划(201410637005); 重庆市基础与前沿研究计划项目(cstc2015jcyjA80020, cstc2015jcyjA80014)

作者简介: 潘秋玲, 女, 1991年9月生, 重庆忠县人, 硕士, 研究方向为病原微生物的防控, E-mail: pql19910907@126.com

* 通讯作者 Corresponding author, E-mail: xujinshan2003@aliyun.com

收稿日期 Received: 2015-06-16; 接受日期 Accepted: 2015-09-29

pombe can regulate heterochromatin assembly and maintain heterochromatin silencing (Bühler and Moazed, 2007). In *Entamoeba histolytica*, 17 candidate miRNAs have been screened and identified by using computational methods (De *et al.*, 2006). Microsporidia were obligated intracellular parasites invading vertebrates and invertebrates. Molecular evidence proved that microsporidia were closely related to fungi (Lee *et al.*, 2008). Up to date, at least twenty-one genomes of microsporidia species have been sequenced (Parisot *et al.*, 2014; Desjardins *et al.*, 2015; Pombert *et al.*, 2015). However, no thorough survey of small RNAs in any of these species mentioned above was reported. *Nosema bombycis*, a typical species of microsporidia, can infect the silkworm (*Bombyx mori*) through vertical transmission from the mother host to their progenitive eggs and horizontal transmission via mouth, and consequently, has a devastating impact on the silkworm industry (Ma *et al.*, 2013; Zhao *et al.*, 2015). The genome of *N. bombycis* is 15.3 Mbp in size, 40% of which composes of the repetitive elements (Pan *et al.*, 2013). However, the defense mechanism against the amplification of repetitive elements remains unclear in the *N. bombycis* genome. It had been reported that small RNAs could defense against transposons in the silkworm (Kawaoka *et al.*, 2008). Therefore, whether small RNAs are associated with transposon defense in *N. bombycis* would be worth excavating.

In the study, we conducted the survey of *N. bombycis* small RNAs to identify the small RNAs associated with transposons and the potential miRNAs in *N. bombycis*, which provides a basis for further study of the biological function of *N. bombycis* small RNAs.

2 MATERIALS AND METHODS

2.1 Test insect

B. mori strain Dazao reserved in our laboratory was cultivated as the 3rd instar larvae after eggs hatched, and then the 3rd instar larvae were infected by spores of *N. bombycis* CQ1, which were isolated from infected silkworms in Chongqing, China, conserved in China Veterinary Culture Collection Center (CVCC number: 102059). At 4 d post infection, some tissues were collected to assay the condition of spores propagation through observation under microscope. Finally, whole bodies of day-3 5th instar larvae were grinded and the total number of spores was counted by using hemocytometer.

2.2 Genomic data resources

Genomic sequences and gene annotation of *N.*

bombycis CQ1 were downloaded from public Silkworm Pathogen Database (SilkPathDB, <http://silkpathdb.swu.edu.cn/silkpathdb/ftpserver>) and GenBank (accession no. ACJZ01000001 – ACJZ01003558).

2.3 Construction of small RNA cDNA libraries for deep sequencing

Whole bodies of day-3 5th instar larvae of *B. mori* infected with total 10^9 spores of *N. bombycis* CQ1 were collected for RNA isolation. After purification using Trizol, total RNAs were immediately stored in ethanol and at -80°C for further use. The total RNA was separated by 15% polyacrylamide gel electrophoresis, and recovered about 16 – 30 nt small RNAs. The 5' RNA adapter (5'-GUUCAGAGUUCUACAGUCCGACGAUC-3') was ligated to the RNAs pool with T4 RNA ligase. The 3' RNA adapter (5'-pUCGUAUGCCGUCUUCUGCUU GidT-3') was subsequently ligated to precipitated RNAs using T4 RNA ligase. Small RNAs ligated with adaptors were then subjected to RT-PCR to produce sequencing libraries. PCR products were purified and small RNA libraries were sequenced using Solexa, a massively parallel sequencing technology. After removing adaptor sequences and low quality ($Q < 30$) sequences, the clean reads were reserved.

2.4 Classification and functional annotation of *N. bombycis* small RNAs

Raw clean small RNA reads were aligned with *N. bombycis* genomic sequences by Bowtie (mismatch ≤ 2 bp) software. Only those matched to genomic sequence were retrieved and called as *N. bombycis* small RNAs. If small RNAs gave hits for repetitive sequence of *N. bombycis* at least once, these sRNAs were categorized as repeat-associated small interfering RNAs (rasiRNAs). If sequences gave hits for non-coding sequence region, they were classified as ncRNAs. Likewise, those that gave hits for protein-coding regions were classified as 'gene'. Others that matched the rDNAs or tRNAs sequence were classified as 'rDNAs' or 'tRNAs', respectively. Those ncRNAs of *N. bombycis* were picked up to predict potential miRNAs by MIREAP (<https://sourceforge.net/projects/mireap/>). Secondary RNA structures of candidate miRNAs were visualized using RNAfold (Hofacker and Stadler, 2006). Three miRNA databases, including MiRAlign (http://bioinfo.au.tsinghua.edu.cn/miralign/miralign_dscems.htm), MiRbase (<http://microrna.sanger.ac.uk/sequences/>) and Rfam (<http://rfam.sanger.ac.uk/>), were retrieved for homologous search of potential miRNAs identified in this study. To intuitively observe homologous sequences of *N.*

bombycis candidate miRNAs in other related microsporidia genomes, both mature miRNAs and precursor miRNAs from *N. bombycis* were Blast-aligned with the public genomic sequences of six microsporidia species including *Nosema antheraeae*, *Nosema ceranae*, *Encephalitozoon cuniculi*, *Encephalitozoon intestinalis*, *Enterocytozoon bieneusi* and *Antonospora locustae*. The parameters for the BLAST search were set as following; matching region covered more than 60% of whole query sequence and pairwise identity thresholds were 80% at nucleic acid level. Subsequently, statistical analysis was performed to investigate the evolutionary conservation of these candidate miRNAs.

3 RESULTS

3.1 Sequencing and classification of small RNAs from *N. bombycis*

A total of 3 844 786 raw reads were obtained by Solexa sequencing. After trimming low-quality, poly (N) adapter and adapter sequences, 2 367 321 clean reads of small RNAs remained. Through matching to *N. bombycis* genomic sequences, 661 609 reads were collected as small RNAs sequences of *N. bombycis* (Table 1). It was notable that there were still 1 705 712 unclassified reads expect *N. bombycis* small RNAs sequences, and these reads should be originated from host domestic silkworm (*B. mori*), because the mixed RNAs isolated from midguts of the

domestic silkworm infected by *N. bombycis* must contain massive RNAs of the silkworm in this study. *N. bombycis* small RNAs showed a unimodal length distribution, with a peak at 25 nt (Fig. 1: A). Seventy percent of cloned *N. bombycis* small RNAs contained a 5' uridine residue (Fig. 1: B), indicative of bases bias in the 5' terminus.

Small RNAs sequences obtained were classified by matching *N. bombycis* genomic annotated database (Table 1). Of the total 661 609 small RNAs, 460 916 reads (75.93%) matched to the non-coding regions of the genome (ncRNA), 162 691 reads (21.68%) matched to the repetitive sequences of the genome (rasiRNA), and 1.73%, 0.65% and 0.01% of small RNAs matched to the protein-coding regions, tRNAs and rDNAs, respectively. All these results indicated that ncRNAs and rasiRNAs were mostly predominant in all *N. bombycis* small RNAs sequences.

Table 1 Sequencing data of *Nosema bombycis* small RNAs

	Number of reads	Total number of nucleic acids
Raw reads	3 844 786	13 567 510
Low quality (< Q30)	1 418 249	49 638 715
Poly(N) reads	7 870	275 450
Clean reads (adapter trimming)	2 367 321	61 598 363
Matching <i>N. bombycis</i>	661 609	15 997 478
Unmatched	1 705 712	45 600 885

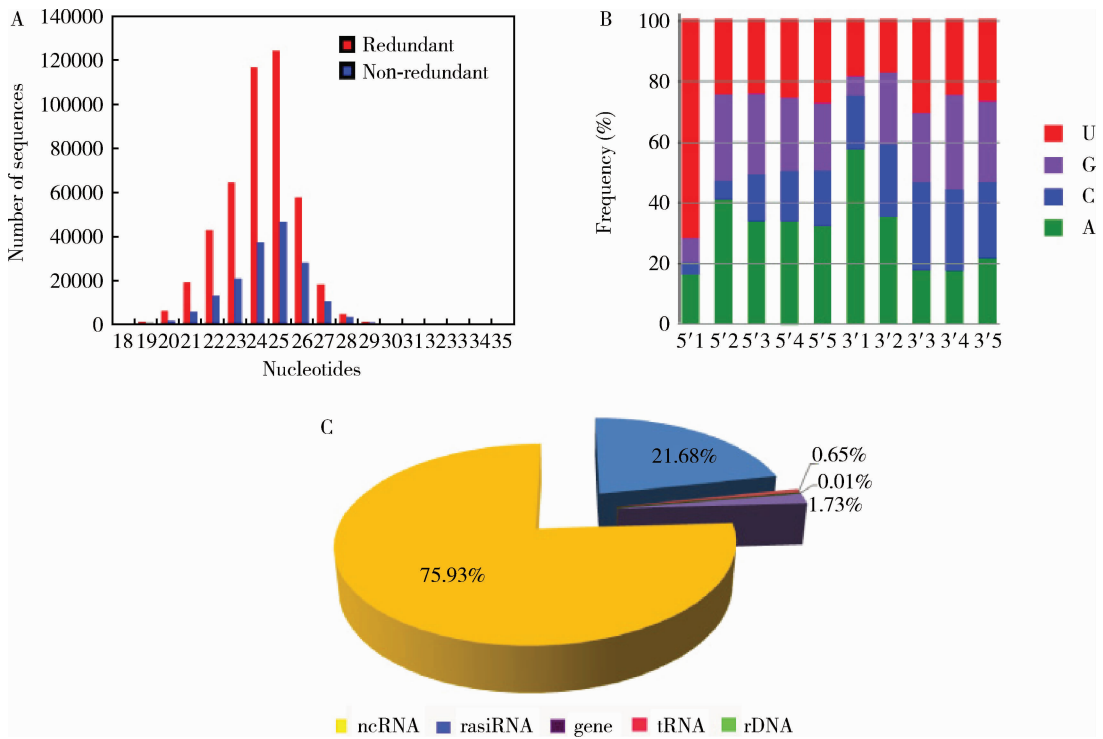


Fig. 1 Characterization of *Nosema bombycis* small RNAs

A: Length distributions of sequenced small RNA reads; B: Nucleotide compositions of sequenced small RNAs; C: Proportions of annotated small RNAs in *N. bombycis*. 5'1 – 5: 1st – 5th bases at the 5' terminus; 3'1 – 5: 1st – 5th bases at the 3' terminus.

3.2 Small RNAs matched with transposons of *N. bombycis*

Most of *N. bombycis* rasiRNAs were found to be associated with well annotated *N. bombycis* transposable elements, including LTR, non-LTR and DNA types. rasiRNAs corresponding to various types of annotated transposons were summarized in Table 2. Of the total 162 691 rasiRNAs, 8 417 rasiRNAs matched with LTR retrotransposons and 6 516 rasiRNAs matched with DNA transposons. Interestingly, rasiRNAs gave hits for all the annotated *N. bombycis* LTR retrotransposons except for the Nbr3 and Nbr8 families. Generally, *N. bombycis* rasiRNAs gave more frequent hits toward antisense orientation of transposons than those hits toward sense orientation (Table 2). Two full-length LTR retrotransposons (Nbr4 and Nbr7), as well as

one DNA transposon (NbTc1), were then chosen to analyze the distribution of sRNAs in the complete elements. After plotting all of the rasiRNAs that were mapped to the targets, the characteristic phasing patterns were identified as in Fig. 2. rasiRNAs are distributed inconsecutively in the three complete elements and give much more frequent hits for antisense orientation than for sense orientation. In the case of Nbr4, rasiRNA is distributed throughout the element, and two peaks of antisense rasiRNAs were found at the sites of 1 200 and 4 100 nt, respectively (Fig. 2; A). In the case of Nbr7, rasiRNA mapped all the regions excluding the two termini, and a peak of antisense rasiRNAs was present at site of 1 300 nt (Fig. 2; B). For NbTc1, sense and antisense rasiRNAs are mainly distributed in the middle of the element (Fig. 2; C).

Table 2 Number and percent of small RNAs corresponding to the different types of annotated transposons in *Nosema bombycis*

Type	Family	Sense	Antisense	Total	Percent	References
DNA	MITE	1 561	1 012	2 573	14.8	He <i>et al.</i> , 2015
	Tc1/mariner	750	2 194	2 944	16.9	Xu <i>et al.</i> , 2010
	Tiger-like	305	694	999	5.7	Pan <i>et al.</i> , 2013
LTR	Nbr1-Nbr8	869	1 663	2 532	14.5	Xu <i>et al.</i> , 2006
	Nbr9-Nbr14	636	1 520	2 156	12.4	Xiang <i>et al.</i> , 2010
	Others	1 942	1 787	3 729	21.4	Pan <i>et al.</i> , 2013
Non-LTR	Nbr4	576	1 914	2 490	14.3	Liu <i>et al.</i> , 2014
Total		6 639	10 784	17 423		

3.3 Identification of *N. bombycis* potential miRNAs

Since 76% of small RNA sequences were ncRNA sequences, we intended to explore whether there were potential miRNAs derived from ncRNAs in *N. bombycis*. After all the ncRNAs were mapped to genome loci, only those 300 bp genomic regions harboring more than 50 ncRNA tags, were chosen as the potential sites to produce miRNA candidate. More than 1 000 sites were picked up to predict miRNAs through MIREAP software under the criterion of energy ≤ -20 kal/mol. And then ‘precursor miRNAs’ containing intricate stem-loop structures were discarded manually by means of RNAfold (Hofacker and Stadler, 2006). Ultimately, 31 candidate miRNAs and miRNA precursors were generated from the *N. bombycis* genome loci sequences. The characteristics of these 31 potential miRNAs were shown in Table 3, and several cases of miRNA stem-loop structure were visualized in Fig. 3(A). While 3 kb interval was defined as the threshold value between two miRNAs (Altuvia *et al.*, 2005), several candidate miRNA genes identified here were arrayed in cluster in the *N. bombycis* genome. Two miRNAs, named as NB-

m0003-5p and NB-m0004-5p, were clustered together since there was only 530 bp interval between them (Fig. 3; B). In order to investigate the presence and conservation of potential miRNAs in other microsporidium species, 31 mature miRNAs and precursor miRNAs from *N. bombycis* were Blast-aligned with the public genomic data of six microsporidia species, which were *N. antheraeae*, *N. ceranae*, *E. cuniculi*, *E. intestinalis*, *E. bieneusi* and *A. locustae*, respectively. As shown in Fig. 3(C), it was found that not only 14 mature miRNAs but also five precursor miRNAs from *N. bombycis* share sequence similarity in *N. antheraeae*. However, only three homologs of *N. bombycis* mature miRNAs exist in *N. ceranae* or *E. bieneusi*, and no hit was found in *E. cuniculi*, *E. intestinalis* and *A. locustae*. These results implied that miRNAs might exist in the common ancestor of microsporidia and undergo the loss events during microsporidia genomic evolution. Prediction of miRNA targets were conducted tentatively through alignment of miRNAs with the 3’ untranslated region (UTR) of any *N. bombycis* genes. As a result, one candidate miRNA named as NB-m0189-3p matched perfectly with the

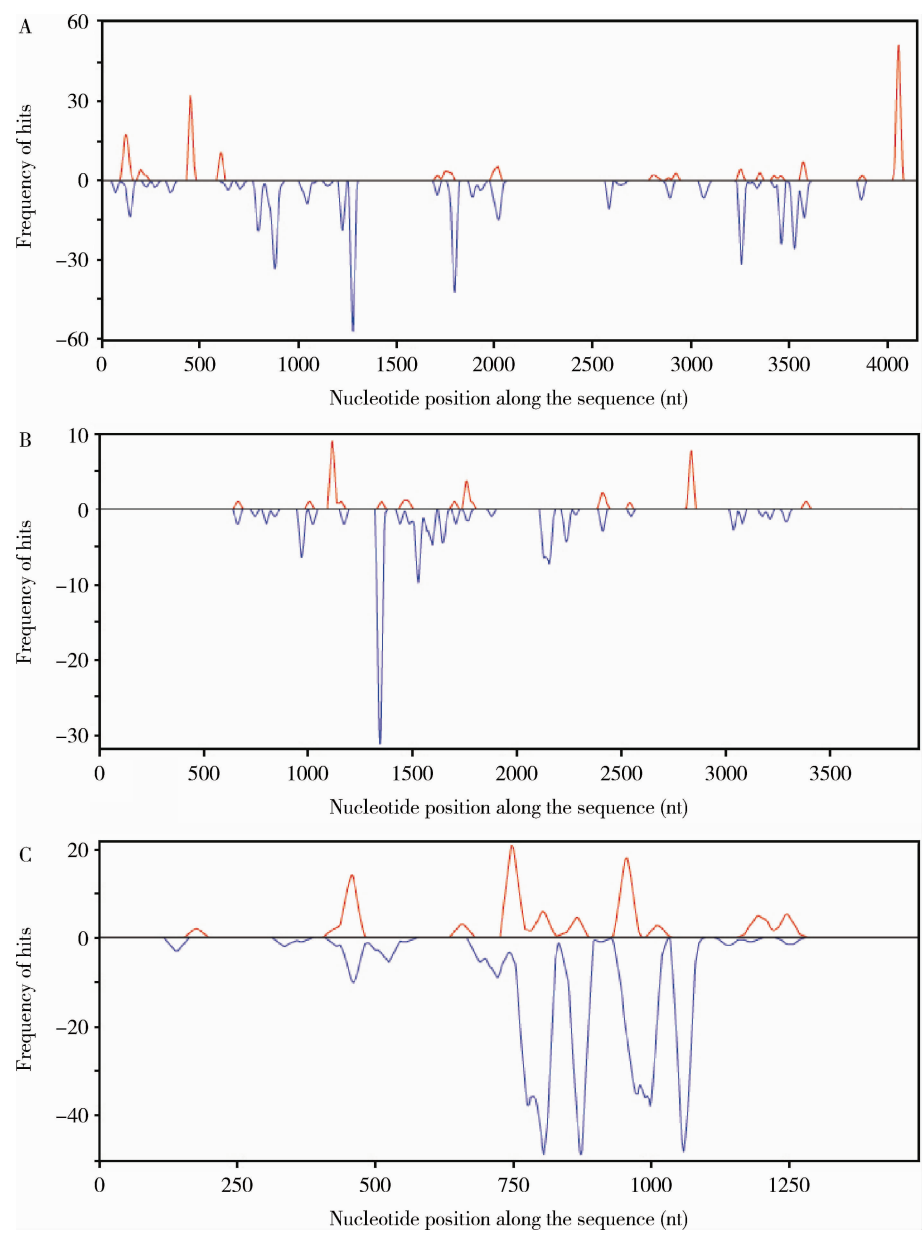


Fig. 2 Density plot of rasiRNAs assigned to Nbr3 (A) , Nbr7 (B) and NbTc1 (C)
Red curve: Sense; Blue curve: Antisense.

3'-UTR of NBO-10g0036 gene (Fig. 3: D).

4 DISCUSSION

Although the key genes responsible for RNAi mechanism have been found recently in microsporidia species (Paldi *et al.*, 2010; Heinz *et al.*, 2012; Wang *et al.*, 2015), there was no report about the small RNAs in microsporidia. In this study, we performed a systematic and integrative analysis of small RNAs in the microsporidia typical species, silkworm parasite *N. bombycis*. After Solexa sequencing and classification of *N. bombycis* small RNAs, 75.93% and 21.68% were found to be ncRNAs and rasiRNAs, respectively. Notably, strong U bias at 5' ends of *N. bombycis* small RNAs was displayed. This phenomenon was therein

consistent with those small RNAs reported in the domestic silkworm (Kawaoka *et al.*, 2008). Our results showed that lots of rasiRNAs gave hits to most well-annotated transposons and mapped their targets in differential manners, indicating that they should have the role in regulating activity of transposon. Moreover, the strong strand bias of *N. bombycis* rasiRNAs toward antisense orientation was usually observed in our study, which might be due to nucleotide constitution of both sense and antisense strands. Previous studies reported that more than 20% of *N. bombycis* genome sequences were comprised of transposable elements, which have an important impact on genomic and gene organization (Xu *et al.*, 2010; Pan *et al.*, 2013), and several transposons still have transcribed activity in *N. bombycis*

Table 3 Characteristics of the predicted candidate miRNAs in *Nosema bombycis*

Locus ID	Length	Scaffold	Start locus	End locus	Scaffold strand	Energy	miR-strand*	MiRAlign	Rfam	miRBase	sRNA sequence	Length (nt)	Number of reads
NB-m0003	148	NBO_1067	189	336	-	-43.0	fwd	+	-	-	TCAGAAACTGTCAGGGGCTGGGCAC	25	3
NB-m0004	81	NBO_1067	866	946	-	-24.3	fwd	+	-	-	TATGCTCGAAGCAAAAGTTGCCAT	25	10
NB-m0040	118	NBO_12	299 145	299 262	+	-38.5	fwd	+	+	-	TGAGGAAGATACTATGCGAAAAGTGC	26	6
NB-m0073	116	NBO_130	16 241	16 356	+	-20.3	fwd	-	-	+	TAGGAATGAGACCATGGGTATAT	23	5
NB-m0084	76	NBO_1385	337	412	-	-26.1	rev	-	-	-	TGTAAAGTTGTGCTGTGAAA	22	8
NB-m0136	158	NBO_14	67 997	68 154	+	-22.1	fwd	-	-	-	TTTACTAAAATAAATCATCTCGGA	25	9
NB-m0172	127	NBO_15	16 297	16 423	+	-37.0	fwd	-	-	-	TTAAATATCGAAAGTGGCAGCGGCAC	27	6
NB-m0189	167	NBO_18	83 567	83 733	-	-36.5	fwd	-	-	-	TAACTGTTTAATGATAGAAAGCGTT	25	4
NB-m0223	134	NBO_21	2 372	2 505	-	-22.9	rev	-	-	+	TACTAACTGTACTTTTCTGGACAC	24	8
NB-m0254	66	NBO_25	142 648	142 713	+	-24.9	fwd	-	-	-	CCTATTACGGGTTGATCGTTC	22	10
NB-m0263	163	NBO_25	149 240	149 402	-	-43.3	amb	-	-	-	TAGGATTATTGAAGAGCGCC	21	17
NB-m0271	105	NBO_28	70 477	70 581	+	-21.5	fwd	-	+	+	GAGATGGCTAAAGTACAGGTAA	21	9
NB-m0274	128	NBO_28	67 289	67 416	-	-29.2	fwd	+	-	-	TAGCACTGTTATGACATGGTCTTC	24	7
NB-m0275	104	NBO_28	68 564	68 667	-	-25.0	fwd	-	-	-	TTTTTGATTAATCCTAAACGGGCAA	25	5
NB-m0288	84	NBO_2	117 692	117 775	-	-26.9	rev	-	-	-	TCGCGATCAGCAGGCGTTCAATGT	24	6
NB-m0307	111	NBO_356	15 989	16 099	+	-24.2	fwd	+	+	-	TGAAAAAGCTTCAGAACCACTGTAC	25	5
NB-m0337	132	NBO_387	6 975	7 106	-	-36.9	rev	+	-	-	TAGCTTTACTGACACATATGCTTCA	25	8
NB-m0390	116	NBO_4	127 036	127 151	+	-26.0	fwd	-	-	-	TAGGTACTAATGGTGTCTTAGCACC	25	5
NB-m0404	100	NBO_509	6 226	6 325	-	-21.4	fwd	-	-	-	TAACAGAGCATATCGACGAGATCAA	25	7
NB-m0407	108	NBO_50	32 345	32 452	+	-25.1	fwd	+	+	-	TATGTAGAACCAAAACACGCCGAA	24	14
NB-m0424	113	NBO_529	3 965	4 077	-	-32.6	rev	+	-	-	TCGTTGATAGATACAGAAATGCCCAC	24	5
NB-m0446	145	NBO_578	1 615	1 759	+	-76.0	fwd	-	-	-	TGAGGCTGTTTCGCTGCTTCTAT	24	6
NB-m0450	99	NBO_58	39 793	39 891	-	-29.9	rev	-	-	-	AAAACTGGCTTAAAGCTACGGCGT	24	8
NB-m0484	111	NBO_6	286 770	286 880	+	-24.2	fwd	+	+	-	TGAAAAAGCTTCAGAACCACTGTAC	25	5
NB-m0485	131	NBO_6	315 909	316 039	+	-34.1	fwd	-	-	-	TGTTGACGAGATTGATTGGATAG	23	37
NB-m0493	108	NBO_6	551 699	551 806	+	-23.2	fwd	+	+	-	TATGTAGAACCAAAACACGCCGAA	24	14
NB-m0500	104	NBO_6	337 152	337 255	-	-21.8	rev	-	-	-	TGGTCAAGACATTTATGCGACAGCA	25	6
NB-m0523	148	NBO_783	199	346	-	-27.5	fwd	-	-	-	AACTTCTGTGGCTTAATGCGATCTTT	27	6
NB-m0555	79	NBO_8	227 122	227 200	+	-22.6	fwd	+	-	-	AAGAAGCTGATAAGAATTGGG	22	4
NB-m0578	134	NBO_8	257 150	257 283	-	-22.9	rev	-	-	+	TACTAACTGTACTTTCTGGACAC	24	8
NB-m0600	142	NBO_9	88 252	88 393	-	-28.1	fwd	-	+	-	TGGAGCTATACTTATGGGTCATTTCA	26	5

* Expression chains associated with prediction and annotation (fwd = forward/rev = reverse). If small RNAs matched with both strands and were not specific, small RNAs were annotated as “ambiguous” (amb).

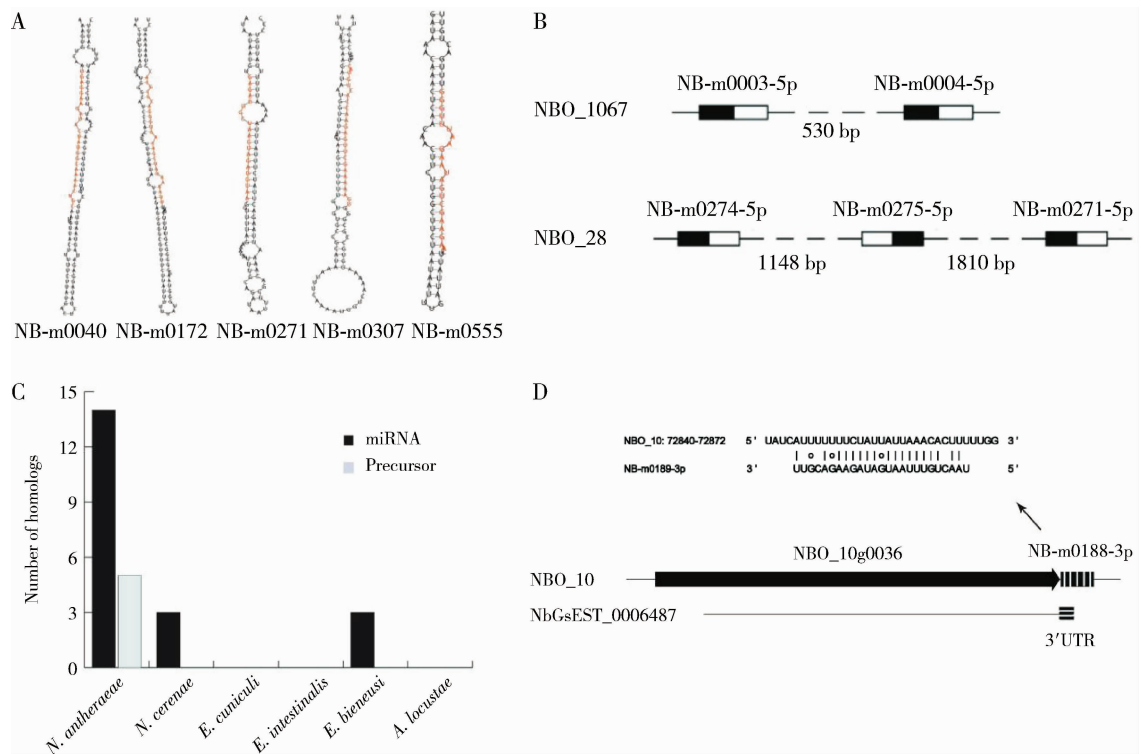


Fig. 3 Characterization of *Nosema bombycis* candidate miRNAs

A: Visualization of secondary structure of four candidate miRNAs from *N. bombycis*; B: miRNA clusters in the genome; C: Homologous sequences of *N. bombycis* miRNAs present in different species; D: Alignment of miRNAs to the 3' UTRs of potential targets.

(Xiang *et al.*, 2010; Xu *et al.*, 2010). Therefore, so many transposons were present in *N. bombycis* genome, it raised the question if there is molecular mechanism of inhibiting the mobilization of transposon and therein keeping the genome stability. Based on our bioinformatic finding in this study, it can imply that these rasiRNAs should be involved in genomic defense system against transposons in *N. bombycis*. However, the main question remains to be unresolved was that how these rasiRNA can silence the transposon in *N. bombycis* genome. Some studies have reported that piRNAs cluster produced from a *Flamenco* locus in X chromosome could mediate the transcriptional level of three retrotransposons gypsy, ZAM and Idefix in *Drosophila melanogaster* (Brennecke *et al.*, 2007), and an interaction between Aubergine-associated antisense piRNAs and AGO3-associated sense piRNAs was regarded to lead to the efficient silencing of transposons in *Drosophila* (Gunawardane *et al.*, 2007). Those findings mentioned above offered the clue to see whether putative AGO gene can interact with rasiRNA and then trigger the silence of active transposon of *N. bombycis*.

In this study, we found several structural features, including hairpin structure, the length of stem-loop, bulge size and thermodynamic stability, which were usually considered as the standards for bioinformatic prediction of miRNAs (De *et al.*, 2006; Chen *et al.*, 2009). On this basis, 31 potential

candidate miRNAs were predicted in *N. bombycis*, some of which were clustered in the certain genomic regions and then might be transcribed as polycistronic structure. In the amphioxus, 45 miRNAs constituted 17 compact clusters, some of which were even conserved in vertebrates (Chen *et al.*, 2009). So our results indicated that the existence of miRNAs clusters might be common in most organisms harboring miRNAs. However, the roles of these candidate miRNAs of *N. bombycis* need to be further resolved. Even if the targeted gene NBO-10g0036 has been predicted by using NB-m0189-3p as query to match perfectly with the 3'-UTR of gene, more evidence is needed to determine whether the NB-m0189-3p can downregulate the transcriptional level of NBO-10g0036 gene based on experiment. Candidate miRNAs share sequence similarity among *N. bombycis*, *N. antheraeae* and *N. ceranae*, indicating that they should be present commonly in *Nosema*. No homolog of miRNAs exists in *E. cuniculi* and *E. intestinalis*, which was consistent with their traits of genomic compaction. However, it was inexplicable that three homologs of miRNA were also found in the *E. bieneusi*, whose genome was considered to be extremely reduced (Akiyoshi *et al.*, 2009). So, more evidence is needed to explain this observation.

All together, this was the first time to identify small RNAs in *N. bombycis*, which should undoubtedly contribute to characterize small RNAs from other

microsporidian species in future. In addition, the findings of *N. bombycis* rasiRNAs and potential miRNAs will help greatly to clarify the origin and function of microsporidia small RNAs.

References

- Akiyoshi DE, Morrison HG, Lei S, Feng X, Zhang Q, Corradi N, Mayanja H, Tumwine JK, Keeling PJ, Weiss LM, Tzipori S, 2009. Genomic survey of the non-cultivable opportunistic human pathogen, *Enterocytozoon bieneusi*. *PLoS Pathogens*, 5(1): e1000261.
- Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ, Tuschl T, Margalit H, 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research*, 33(8): 2697–2706.
- Brenneke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ, 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, 128: 1089–1103.
- Bühler M, Moazed D, 2007. Transcription and RNAi in heterochromatic gene silencing. *Nat. Struct. Mol. Biol.*, 14: 1041–1048.
- Chen X, Li Q, Wang J, Guo X, Jiang X, Ren Z, Weng C, Sun G, Wang X, Liu Y, Ma L, Chen JY, Wang J, Zen K, Zhang J, Zhang CY, 2009. Identification and characterization of novel amphioxus microRNAs by Solexa sequencing. *Genome Biology*, 10: R78.
- De S, Pal D, Ghosh SK, 2006. *Entamoeba histolytica*: computational identification of putative microRNA candidates. *Experimental Parasitology*, 113: 239–243.
- Desjardins CA, Sanscrainte ND, Goldberg JM, Heiman D, Young S, Zeng Q, Madhani HD, Becnel JJ, Cuomo CA, 2015. Contrasting host-pathogen interactions and genome evolution in two generalist and specialist microsporidian pathogens of mosquitoes. *Nature Communications*, 6: 7121.
- Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP, 2009. RNAi in budding yeast. *Science*, 326: 544–550.
- Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G, 2008. A human snoRNA with microRNA-like functions. *Molecular Cell*, 32: 519–528.
- Girard A, Sachidanandam R, Hannon GJ, Carmell MA, 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, 442: 199–202.
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC, 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science*, 315: 1587–1590.
- He Q, Ma Z, Dang X, Xu J, Zhou Z, 2015. Identification, diversity and evolution of MITEs in the genomes of microsporidian *Nosema* parasites. *PLoS ONE*, 10(4): e0123170.
- Heinz E, Williams TA, Nakjang S, Noël CJ, Swan DC, Goldberg AV, Harris SR, Weinmaier T, Markert S, Becher D, Bernhardt J, Dagan T, Hacker C, Lucocq JM, Schweder T, Rattei T, Hall N, Hirt RP, Embley TM, 2012. The genome of the obligate intracellular parasite *Trachipleistophora hominis*: new insights into microsporidian genome dynamics and reductive evolution. *PLoS Pathogens*, 8(10): e1002979.
- Höck J, Meister G, 2008. The Argonaute protein family. *Genome Biology*, 9: 210.
- Hofacker IL, Stadler PF, 2006. Memory efficient folding algorithms for circular RNA secondary structures. *Bioinformatics*, 22(10): 1172–1176.
- Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RHA, Hannon GJ, Draper BW, Ketting RF, 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in zebrafish. *Cell*, 129: 69–82.
- Kawaoka S, Hayashi N, Katsuma S, Kishino H, Kohara Y, Mita K, Shimada T, 2008. *Bombyx* small RNAs: genomic defense system against transposons in the silkworm. *Bombyx mori*. *Insect Biochem. Molec. Biol.*, 38: 1058–1065.
- Kim VN, Han J, Siomi MC, 2009. Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology*, 10: 126–137.
- Lau NC, Robine N, Martin R, Chung WJ, Niki Y, Berezikov E, Lai EC, 2009. Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Research*, 19: 1776–1785.
- Lee SC, Corradi N, Byrnes EJ, Torres-Martinez S, Dietrich FS, Keeling PJ, Heitman J, 2008. Microsporidia evolved from ancestral sexual fungi. *Curr. Biol.*, 18(21): 1675–1679.
- Liu LL, Wu CM, Zhang XY, Xu JS, Zhou ZY, 2014. Identification and phylogenetic analysis of LINE retrotransposons R4 in *Nosema bombycis* and *Nosema pernyi*. *Science of Sericulture*, 40(4): 681–687. [刘兰兰, 吴春漫, 张小燕, 许金山, 周泽扬, 2014. 家蚕微孢子虫和柞蚕微孢子虫 LINE 类反转录转座子 R4 的鉴定及系统进化分析. *蚕业科学*, 40(4): 681–687]
- Ma Z, Li C, Pan G, Li Z, Han B, Xu J, Lan X, Chen J, Yang D, Chen Q, Sang Q, Ji X, Li T, Long M, Zhou Z, 2013. Genome-wide transcriptional response of silkworm (*Bombyx mori*) to infection by the microsporidian *Nosema bombycis*. *PLoS ONE*, 8(12): e84137.
- MacRae IJ, Zhou K, Doudna JA, 2007. Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.*, 14(10): 934–940.
- Obbard DJ, Finnegan DJ, 2008. RNA interference: endogenous siRNAs derived from transposable elements. *Curr. Biol.*, 18(13): R561–R563.
- Paldi N, Glick E, Oliva M, Zilberberg Y, Aubin L, Pettis J, Chen Y, Evans JD, 2010. Effective gene silencing in a microsporidian parasite associated with honeybee (*Apis mellifera*) colony declines. *Appl. Environ. Microbiol.*, 76(17): 5960–5964.
- Pan G, Xu J, Li T, Xia Q, Liu SL, Zhang G, Li S, Li C, Liu H, Yang L, Liu T, Zhang X, Wu Z, Fan W, Dang X, Xiang H, Tao M, Li Y, Hu J, Li Z, Lin L, Luo J, Geng L, Wang L, Long M, Wan Y, He N, Zhang Z, Lu C, Keeling PJ, Wang J, Xiang Z, Zhou Z, 2013. Comparative genomics of parasitic silkworm microsporidia reveal an association between genome expansion and host adaptation. *BMC Genomics*, 14: 186.
- Parisot N, Pelin A, Gasc C, Polonais V, Belkorchia A, Panek J, Alaoui HE, Biron DG, Brasset É, Vauray C, Peyret P, Corradi N,

- Peyretailade É, Lerat E, 2014. Microsporidian genomes harbor a diverse array of transposable elements that demonstrate an ancestry of horizontal exchange with metazoans. *Genome Biol. Evol.*, 6(9): 2289–2300.
- Pombert JF, Haag KL, Beidas S, Ebert D, Keeling PJ, 2015. The *Ordospora colligata* genome: evolution of extreme reduction in microsporidia and host-to-parasite horizontal gene transfer. *mBio*, 6(1): e02400-14.
- Saito K, Nishida KM., Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC, 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes & Development*, 20: 2214–2222.
- Sinkkonen L, Hugenschmidt T, Berninger P, Gaidatzis D, Mohn F, Artus-Revel CG, Zavolan M, Svoboda P, Filipowicz W, 2008. MicroRNAs control *de novo* DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.*, 15(3): 259–267.
- Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD, 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, 313: 320–324.
- Wang K, Liang C, Liu J, Xiao H, Huang S, Xu J, Li F, 2014. Prediction of piRNAs using transposon interaction and a support vector machine. *BMC Bioinformatics*, 15: 419.
- Wang L, He CM, Dang XQ, Xu JS, Zhou ZY, 2015. Identification, evolutionary and transcriptional activity analyses of a Dicer-like gene in microsporidium *Nosema bombycis*. *Science of Sericulture*, 41(1): 64–71. [王玲, 何承民, 党晓群, 许金山, 周泽扬, 2015. 家蚕微孢子虫的一个 Dicer 相似蛋白基因的鉴定及系统进化和转录活性分析. 蚕业科学, 41(1): 64–71]
- Xiang H, Pan G, Zhang R, Xu J, Tian L, Li W, Zhou Z, Xiang Z, 2010. Natural selection maintains the transcribed LTR retrotransposons in *Nosema bombycis*. *Journal of Genetics and Genomics*, 37(5): 305–314.
- Xu J, Pan G, Fang L, Li J, Tian X, Li T, Zhou Z, Xiang Z, 2006. The varying microsporidian genome: existence of long-terminal repeat retrotransposon in domesticated silkworm parasite *Nosema bombycis*. *International Journal for Parasitology*, 36: 1049–1056.
- Xu J, Wang M, Zhang X, Tang F, Pan G, Zhou Z, 2010. Identification of NbME MITE families: potential molecular markers in the microsporidia *Nosema bombycis*. *J. Invertebr. Pathol.*, 103(1): 48–52.
- Zhao W, Hao Y, Wang L, Zhou Z, Li Z, 2015. Development of a strategy for the identification of surface proteins in the pathogenic microsporidian *Nosema bombycis*. *Parasitology*, 142: 865–878.

家蚕微孢子虫中小 RNAs 的全基因组鉴定与分析

潘秋玲¹, 李 田², 何 强¹, 马振刚¹, 范晓东¹, 张小燕¹,
王艳丽¹, 周泽扬^{1,2}, 许金山^{1,*}

(1. 重庆师范大学生命科学学院, 重庆 401331; 2. 西南大学家蚕基因组生物学国家重点实验室, 重庆 400716)

摘要:【目的】家蚕 *Bombyx mori* 微孢子病是蚕业生产上的毁灭性病害,家蚕微孢子虫 *Nosema bombycis* 是该病的病原,可经卵垂直传播和经口水平传播。为了探索家蚕微孢子虫中对重复元件的抵御以及对基因转录调控的潜在方式,本研究拟在基因组水平上对该物种的小 RNAs 进行全面系统的分析,鉴定与转座子相关的小 RNAs 和潜在的 miRNAs。【方法】从感染家蚕微孢子虫的家蚕中肠中提取总 RNA,分离小片段 RNA 并反转录后,进行 Solexa 高通量测序。通过生物信息学方法对小 RNAs 进行分类及功能注释,鉴定起源于家蚕微孢子虫不同类型转座子的小 RNAs,并对潜在的 miRNA 进行预测分析。【结果】家蚕微孢子虫小 RNAs 的长度主要是 24 和 25 nt,其中大部分序列表现出 5'末端的尿嘧啶偏好性。家蚕微孢子虫中存在丰富的与转座子相关联的小 RNAs,并且与转座子标准序列匹配的反义小 RNAs 明显多于正义小 RNAs。同时,鉴定获得了 31 个候选 miRNAs,部分为 *Nosema* 属的其他孢子虫中所共有,暗示其在微孢子虫基因组进化上具有保守性。【结论】首次鉴定到家蚕微孢子虫的转座子相关性小 RNAs,暗示小 RNAs 在家蚕微孢子虫基因组对转座子防御过程中起到作用,31 个潜在的 miRNAs 为家蚕微孢子虫 miRNAs 的功能验证提供了后续靶标。

关键词: 家蚕微孢子虫; 小 RNA; 转座子; miRNA; 基因组

中图分类号: Q966 **文献标识码:** A **文章编号:** 0454-6296(2015)11-1213-09